ELECTRONIC **WORKSHOPS IN COMPUTING**

Series edited by Professor C.J. van Rijsbergen

D. J. Duke , University of York, UK and A.S. Evans, University of Bradford, UK (Eds)

2nd BCS-FACS Northern Formal Methods Workshop

Proceedings of the 2nd BCS-FACS Northern Formal Methods Workshop, Ilkley, 14-15 July 1997

The Geometry of Distributions in Formal Methods

M. Mac an Airchinnigh and A.P. Hughes

Published in collaboration with the
British Computer Society

BCS
FACS

Springer

The Geometry of Distributions in Formal Methods

Mícheál Mac an Airchinnigh
Arthur Hughes
University of Dublin,
Trinity College, Dublin, Ireland
e-mail: {mmaa@cs.tcd.ie, mmaa@kmtech.ie}

Abstract

The act of distributing and the resulting distribution are notions which lie at the kernel of any distributed system. The basic algebra of such distributions and their use in formal specifications has already been developed in terms of indexed monoids (i.e., function spaces with valuations in monoids) and their morphisms. Complementary to such algebra is a body of emerging geometry/topology of formal specifications, one critical aspect of which is the fibre bundle, and more generally the sheaf.

Fibre bundles are used to model the nature and shape of geometrical objects and to associate a field with points in a space. They find particular application in theoretical physics, for example. We demonstrate here that fibre bundles occur naturally in specifications and models associated with formal methods.

keywords: distributed system, fibre bundle, formal method, geometry, sheaf.

1 Prologue

In an earlier paper *Towards a New Conceptual Framework for the Modelling of Dynamically Distributed Systems* [18] we set forth a *framework* whereby we might set some recent research results into the formal modelling and specification of distributed systems and, in particular, to the harmonious arrangement whereby process algebras might comfortably settle with the standard model-theoretic methods of Z and VDM.

Of even greater interest and concern was the possibility that there might be a *geometry* of formal methods, a geometry which would *complement* the algebra which had proven so fruitful and successful over a period of twenty years. We were fully aware of the Cartesian duality, *algebra* ↔ *geometry*, thereby entailed. A certain success in this regard was reported in a previous workshop in Ilkley [18] where we demonstrated rigorously that the tail-recursive forms of the *len* homomorphism on words was isomorphic to a translation subgroup of the real number line.

1.1 Geometric algebra

The specific advantage of having a geometry of a specification in addition to the algebra lies in the fact that with the geometry one immediately grasps the whole picture at a glance as it were. The geometry provides a qualitative view of specifications (cf., Poincaré: "Ces théorèmes ont été présentés sous une forme géométrique qui avait à mes yeux l'avantage de mieux faire comprendre la genèse de mes idées ..." quoted in [2, 83]). With the algebra, necessary in its own right for the subsequent development and programming, everything is detail. Since the geometry of a specification bears little resemblance to the 'every day' notion of geometry that most people seem to have, some care is needed in its exposition. This is as true for those in formal methods as for others.

It would appear that algebra and logic is the 'natural language' of representation for a formal specification. That there might be a corresponding geometry is entirely plausible. Since the geometry is brought in to

$$\begin{tikzcd}(\mathcal{P}W, \cup, \emptyset) \arrow[r, shift left, "\mathcal{D}_{\mathcal{S}}"] & (L \rightarrow \mathcal{P}W, \circledcup, \theta) \arrow[l, shift left, "{}^{\cup}/ \circ \mathsf{rng}"]\end{tikzcd}$$

Figure 1: A distribution and its retrieval

elucidate the algebra of the specification we call it **geometric algebra**. Our work is, of course, directly related to *Logical Geometry* or geometric logic [10] [8, 458 *et seq.*] [3], it being the counterpart to the constructionist intuitionistic mathematics that we normally use in formal specifications.

In this paper we will focus exclusively on the single notion of **distribution** and examine it in detail. To motivate the discussion we consider some elementary examples. Let there be given a collection of words, $U \in \mathcal{P}W$, which is used to model a spelling-checker dictionary such as might be used in the trade-marked game SCRABBLE. Let us now consider the dictionary to be distributed over a network (such as that of the World-Wide Web). We shall insist that such a distribution is a **partition** of the original dictionary, i.e., that a word can not be found in more than one location. Such a distribution δ in $L \rightarrow \mathcal{P}W$ is an element of an indexed monoid [16], denoted $((L \rightarrow \mathcal{P}W)', \circledcup, \theta)$, which is essentially a function space with target the base monoid, here $(\mathcal{P}W, \cup, \emptyset)$, the monoid of sets under union. The priming of the space denotes the exclusion of elements of the form $l \mapsto \emptyset$. Composition of distributions δ_j and δ_k is 'pointwise'

$$(\delta_j \circledcup \delta_k)(l) = \delta_j(l) \cup \delta_k(l)$$

For the present we will denote the act of distributing thus

$$\mathcal{D}_{\mathcal{S}} \colon \mathcal{P}W \longrightarrow (L \rightarrow \mathcal{P}W)$$

and shall explore the meaning to be given to $\mathcal{D}_{\mathcal{S}}(U) = \delta$ in a geometrical setting, where U denotes a set of words and $S = \mathsf{dom}\,\delta$ denotes the set of sites or locations.

To *collect together* the distributed components of the dictionary is usually known as a retrieval and one strives to ensure that such a retrieval is a homomorphism. Such a collecting together is the *inverse operation* of the act of distributing and the corresponding structures or spaces are complements. In the case of the distributed dictionary the retrieval homomorphism is a reduction with respect to set union, ${}^{\cup}/ \circ \mathsf{rng}$, as shown in Fig. 1. Thus for distributed dictionaries δ_i and δ_j we have

$${}^{\cup}/ \circ \mathsf{rng}(\delta_i \circledcup \delta_j) = {}^{\cup}/ \circ \mathsf{rng}(\delta_i) \cup {}^{\cup}/ \circ \mathsf{rng}(\delta_j)$$

The *creative* aspect of such a dictionary is the addition of a new word. Let A_u denote the addition of a new word u to the original dictionary and $A_{l,u}$ the addition of the same word to the distributed dictionary at location l. We shall not worry for the moment about the **mechanism** which determines the location at which the new word is added to the distributed dictionary. The specification of each of the operations is

$$A_u\delta = \delta \cup \{u\}$$

and

$$A_{l,u}\delta = \delta \circledcup [l \mapsto \{u\}]$$

Repeated addition of new words gives rise to the monoids $(\mathcal{P}W, \cup, \emptyset)$ and $(L \rightarrow \mathcal{P}W, \circledcup, \theta)$, respectively.

From the perspective of algebra and formal specifications it is important to note that not everything in $L \rightarrow \mathcal{P}W$ will denote a distributed dictionary. Therefore, we must supply an appropriate well-formedness constraint or invariant which singles out that space of distributed dictionaries in which we are interested. In addition, the specifications of the addition operations are incomplete. Both require pre-conditions to single out the appropriate meanings. In the case of $A_{l,u}$ we must unfold the operation into the two basic ones of extension and strict override:

$$A_{l,u}\delta = \delta \circledcup [l \mapsto \{u\}] = \begin{cases} \delta \sqcup [l \mapsto \{u\}], & \text{if } \neg\chi_\delta(l) \\ \delta \dagger [l \mapsto \delta(l) \cup \{u\}], & \text{otherwise} \end{cases}$$

We will show that in the geometry of such distributions there is no need for invariants and pre-conditions. The geometry will say it all. On the other hand the geometry is complementary to the algebra. Even if the geometry is determined one will still have to return to the algebra together with the invariants and pre-conditions.

2 The algebra of distributions

We have already introduced a simple example of a distribution—the spelling checker dictionary. Now we wish to look at another well-known and practical distribution—the hash table. Our primary exemplar or prototype of a **distribution function** in computing is the hash function which is used to map words to distinct addresses or locations. In the case that two distinct words hash to the same address we use an overflow list. This implementation strategy is usually known as *hashing with overflow chaining on collision* and the resulting storage structure a hash table [13, 513]. We will wish to extend this prototype to cover a variety of distributions and, therefore, we need to abstract away the concrete details. A rudimentary specification of such a hash table is given in the appendix. From the specification it is clear that everything depends on the nature of the hash function.

Let W denote the space of words, which in general we shall take to stand for the space of words over the alphabet Σ, denoted by $\Sigma^\star$, or the corresponding free group over Σ, denoted by $FG(\Sigma)$. The context will make clear which interpretation is more appropriate. Let $\mathcal{P}W$ denote the space of all sets of words taken from W, the powerset domain of W, which in topological terms is the discrete topology for W in which every subset (or element of $\mathcal{P}W$) is open (see Appendix). The additive abelian group of integers modulo a prime p is denoted by $\mathbf{Z}_p$ or alternatively by the quotient $\mathbf{Z}/p\mathbf{Z}$. The space of partial maps from W to $\mathbf{Z}_p$ is the usual domain of hash functions and the subspace of total functions from W to $\mathbf{Z}_p$ is denoted by $\mathbf{Z}_p^W$.

We have already demonstrated clearly that whereas we work within constructive mathematics there is a seamless transition from the classical mathematics of total functions to the discrete mathematics of finite partial maps [18]. Consequently, let (η, U) in $\mathbf{Z}_p^W \times \mathcal{P}W$ denote a *total* function η which distributes elements of the set U into distinct locations. The resulting distribution may be expressed by the restriction of the total function η to the set of elements U, expressed thus $\lhd_U(\eta)$ (or classically as $\eta|_U$). A hash table is often expressed as the inverse map, $(\lhd_U(\eta))^{-1}$, and expressed in the "completed" or totalized form as $\emptyset^{\mathbf{Z}_p} \dagger (\lhd_U(\eta))^{-1}$. A simple example will make this clear.

Let $U = \{u, v, w\}$ be a set of words which hash under η to the not necessarily distinct elements $\eta(u) = i$, $\eta(v) = j$ and $\eta(w) = k$ in $\mathbf{Z}_p$. Then the distribution is given by

$$\underset{U}{\lhd}(\eta) = \begin{bmatrix} u & \mapsto & i \\ v & \mapsto & j \\ w & \mapsto & k \end{bmatrix}$$

Distributions are combined using *glueing*:

$$\underset{U}{\lhd}(\eta) \cup \underset{V}{\lhd}(\eta) = \underset{U \cup V}{\lhd}(\eta)$$

We will find that this combination is in full agreement with the geometric form. Supposing that words u and v hash to the same location under η, i.e., that $i = j$, then the hash table has the usual form

$$\mathcal{D}_{\eta,\mathcal{S}}(U) = (\underset{U}{\lhd}(\eta))^{-1} = \begin{bmatrix} i & \mapsto & \{u, v\} \\ k & \mapsto & \{w\} \end{bmatrix}$$

It is this latter form which suggested to us the idea of a hash table, and distributions in general, as a fibre bundle, discussed later in a separate section on the geometry of specifications. However, looking ahead, we wish to emphasize here that this form does **not** prove to be very fruitful. Note also that we have chosen to use $\mathcal{D}_{\eta,\mathcal{S}}$ to denote the resulting distribution and that we have

$$({}^{\cup}/ \circ \mathsf{rng}) \circ \mathcal{D}_{\eta,\mathcal{S}} = \mathcal{I}$$

Before turning to consider in detail the geometry of distributions it will be helpful and useful to explore further some of the algebra which is now suggested by the problem of distribution.

2.1 Distributions with algebraic structure

Specifically, it is of interest to pose the question whether or not such distributions (cf., [11]) can be combined and in analogy to Fourier analysis, we seek a mechanism by which a suitable distribution function η might be obtained as a linear combination of *basis* functions η_i:

$$\eta(w) = (a_1\eta_1(w) + a_2\eta_2(w) + \ldots + a_n\eta_n(w)) \pmod{p}$$

where the coefficients a_j are taken from the (finite) field $\mathbf{Z}_p$. As a first step in this direction we consider the composition

$$(\eta, S) + (\kappa, T) = (\eta + \kappa, S \cup T)$$

where the sum of functions η and κ is given pointwise:

$$(\eta + \kappa)(w) = \eta(w) + \kappa(w) \pmod{p}$$

Choosing $\kappa = \eta$ gives us $(\eta, S) + (\eta, T) = (\eta + \eta, S \cup T)$ and, in general, $\eta + \eta \neq \eta$. This result is at first disconcerting. In the concrete case of hashing, we expect that if η is a hash function and it we hash a set S and then hash a set T under the same function η then we expect the result to be exactly the same as if we hashed $S \cup T$ under η. Moreover, if we hash a single word, say w, i.e., if $S = T = \{w\}$, then $(\eta+\eta)(w) \neq \eta(w)$ implies that if we hash it twice in a row we will not get the same location!

On the other hand we can introduce constant functions η_i which map a strict subset S_i of words onto i. In this case, using map extension, we may form the hash function

$$\eta = \eta_0 \sqcup \eta_1 \sqcup \ldots \eta_i \sqcup \ldots \sqcup \eta_{p-1}$$

such that the corresponding *disjoint* domains add up to the original word space, i.e., $S_0 \cup S_1 \cup \ldots \cup S_i \cup \ldots \cup S_{p-1} = W$. If we are prepared to allow for the possibility of some or all of the S_i to overlap, then we may introduce *glueable hash functions* η_i (i.e., functions which agree on their common domain intersections) which are no longer constant but which combine in the expected manner:

$$\eta = \eta_0 \cup \eta_1 \cup \ldots \eta_i \cup \ldots \cup \eta_{p-1}$$

Without wishing to prejudice the outcome, let us for the present *abandon the interpretation in the domain of hashing* and explore the consequences of the law of composition. In other words, we anticipate that in the more general setting of distributions, it may very well be the case that the law of composition is fruitful. First, it may readily be demonstrated that the constant $(0^W, \emptyset)$ is an identity element under addition:

$$(0^W, \emptyset) + (\eta, S) = (\eta, S) = (\eta, S) + (0^W, \emptyset)$$

The operation is associative:

$$(\eta, S) + \Big((\kappa, T) + (\lambda, U)\Big) = \Big((\eta, S) + (\kappa, T)\Big) + (\lambda, U)$$

and commutative

$$(\eta, S) + (\kappa, T) = (\kappa, T) + (\eta, S)$$

Thus our space of distributions is an additive abelian monoid. Might it be possible that we could construct a group of distributions? In short, for a distribution (η, S) we seek an inverse distribution (κ, T) such that

$$(\eta, S) + (\kappa, T) = (0^W, \emptyset)$$

By the law of composition we have $\eta + \kappa = 0^W$ and $S \cup T = \emptyset$. Since $\mathbf{Z}_p$ is an additive abelian group then there always is a $\kappa(w)$ such that $\eta(w) + \kappa(w) \pmod p = 0^W(w) = 0$. Hence there is a total function κ such $\eta + \kappa = 0^W$. For the second part it is clear that there is no T such that $S \cup T = \emptyset$. However, were one to use symmetric difference, $\triangle$, then we do have an abelian group on sets.

> **Definition 2.1** *The* symmetric difference *of two sets S and T, denoted $S \triangle T$, is the union of the two sets less their common intersection. Traditionally, this is written as*
>
> $$S \triangle T = (S \cup T) \backslash (S \cap T) = (S \backslash T) \cup (T \backslash S)$$
>
> *and is equivalent to*
>
> $$S \triangle T = \underset{S \cap T}{⩤} (S \cup T) = \underset{T}{⩤} (S) \cup \underset{S}{⩤} (T)$$

Consequently, with the inner law of composition defined by

$$(\eta, S) + (\kappa, T) = (\eta + \kappa, S \triangle T)$$

then $(\mathbf{Z}_p^W \times \mathcal{P}W, +)$ is an abelian group of distributions.

2.2 Construction

Let us return now to the domain of hashing. The algebraic excursion above did not seem to advance us much on the path to modelling the real hash table. A more conventional approach is to consider the construction of a hash table through the addition of a new word which may readily be modelled by

$$A_w(\eta, S) = (\eta, S \cup \{w\})$$

and since one normally derives the inner law of composition on structures from such additive operations

$$\begin{aligned}(A_u \circ A_v)(\eta, S) &= A_u(A_v(\eta, S)) \\ &= A_u(\eta, S \cup \{v\}) \\ &= (\eta, S \cup \{v\} \cup \{u\}) \\ &= A_{\{u,\, v\}}(\eta, S)\end{aligned}$$

then we have the result

$$A_T(\eta, S) = (\eta, S \cup T) = A_{(S \cup T)}(\eta, \emptyset)$$

Therefore, we would expect to find a similar construction in the geometric form. It does indeed turn out to be the case.

2.3 Rehashing

Finally, we will use rehashing as a check on the validity of the geometric form. Rehashing is the practical process which is required whenever a hash table becomes ‘too full’. Formally, we may think of rehashing in terms of a *morphism* α from $\mathbf{Z}_p^W \times \mathcal{P}W$ to $\mathbf{Z}_q^W \times \mathcal{P}W$ where p and q are primes and $q > p$. For complete generality we shall allow also the possibility of $p \geq q$. The map $\alpha\colon (\eta, S) \mapsto (\kappa, S)$ transforms a hash function η into a hash function κ while leaving the set of words hashed invariant. Applying α to the restricted hash function gives

$$\alpha(\underset{S}{\lhd}\,\eta) = \underset{S}{\lhd}\,\alpha(\eta) = \underset{S}{\lhd}\,\kappa$$

Therefore, we expect that rehashing will turn out to be a *morphism* also of suitable geometric objects.

3 Isomorphic Representations

It is remarkable how the choice of representation suggests a development that is totally unexpected. For many years we had used a simple ‘implementation-oriented’ model of a hash table. With the introduction of the more abstract $(\eta, S) \in \mathbf{Z}_p^W \times \mathcal{P}W$ we have been able not only to model the original hash table accurately and concisely but also to develop a more general theory of distributions. We shall see later how this model also opened the way to another ‘chapter’ in the application of geometry of formal methods.

One of the significant advantages of embracing algebra in its totality is the potentiality of transferring to and exploring other isomorphic representations. Here we illustrate the new results arrived at in this process. Using the notation $\mathbf{p} = \{0, \ldots, p-1\}$ (see [6]) and recalling that $\mathcal{P}W$ is isomorphic to $\mathbf{2}^W$ we may express our fundamental model in the form of the product of two total function spaces $(\eta, \chi_S) \in \mathbf{p}^W \times \mathbf{2}^W$, where χ_S denotes the (total) characteristic function associated with S, defined simply as follows:

$$\chi_S(w) = \begin{cases} 1, & \text{if } w \in S \\ 0, & \text{otherwise} \end{cases}$$

Using our example of the previous section we have

$$(\eta, \chi_S) = \left(\begin{bmatrix} \vdots & \mapsto & \vdots \\ u & \mapsto & i \\ v & \mapsto & j \\ w & \mapsto & k \\ \vdots & \mapsto & \vdots \end{bmatrix}, \begin{bmatrix} \vdots & \mapsto & 0 \\ u & \mapsto & 1 \\ v & \mapsto & 1 \\ w & \mapsto & 1 \\ \vdots & \mapsto & 0 \end{bmatrix} \right)$$

where only those words which are hashed map to the value 1 in the characteristic function χ_S.

There is clearly a one to one correspondence between S and χ_S. The advantage of the latter is simply that it is a total function. Since we are in an isomorphic space, then we immediately have a group of distributions as before, $(\mathbf{p}^W \times \mathbf{2}^W, +)$, where the law of composition is

$$(\eta, \chi_S) + (\kappa, \chi_T) = (\eta + \kappa, \chi_S \bigtriangleup \chi_T)$$

and the symmetric difference on characteristic functions is just that on sets. Before we give a formal definition of symmetric difference on characteristic functions, let us consider a simple example. If u and v are two distinct words such that $\eta(u) = i$ and $\eta(v) = j$, then the union of $\{u\}$ and $\{v\}$ is the same as the symmetric difference of $\{u\}$ and $\{v\}$:

$$\{u\} \cup \{v\} = \{u, v\} = \{u\} \bigtriangleup \{v\}$$

On the other hand, we will want the symmetric difference of the corresponding total functions $\chi_{\{u\}}$ and $\chi_{\{v\}}$ to have the form

$$\chi_{\{u\}} \triangle \chi_{\{v\}} = \chi_{\{u, v\}} = \chi_{(\{u\} \triangle \{v\})}$$

Using a totalizer [18], we may write the characteristic function in terms of an override operator on the zero function:

$$\chi_{\{u, v\}} = 0^W \dagger [u \mapsto 1, v \mapsto 1]$$

With this notation we are able to elaborate the meaning of symmetric difference on characteristic functions:

$$\chi_{\{u\}} \triangle \chi_{\{v\}} = (0^W \dagger [u \mapsto 1, v \mapsto 0]) \triangle (0^W \dagger [u \mapsto 0, v \mapsto 1])$$

Note in particular that considering characteristic functions as sets of pairs and applying either set union or symmetric difference leads to **strange** results:

$$\{\dots, (u, 1), (v, 0), \dots\} \cup \{\dots, (u, 0), (v, 1), \dots\} = \{\dots, (u, 0), (u, 1), (v, 0), (v, 1), \dots\}$$

Since there can never be any possibility of confusion then it is quite proper to use $\chi_{\{u, v\}}$ to stand for the right hand side partial function in a totalizer of the form

$$0^W \dagger [u \mapsto 1, v \mapsto 1]$$

This identification of the total function χ_S with the, as yet to be totalized partial function, is of considerable practical importance in formal specification.

Before we complete this section by introducing our final example of an isomorphic representation of a distribution, it seems appropriate to remark here upon how one may generalize from ‘binary’ or ‘boolean’ systems, as expressed by $\mathbf{2} = \{0, 1\}$.

If we should choose $\mathbf{3} = \{0, 1, 2\}$ as the basis for our characteristic function space, $\mathbf{3}^W$, then we have the basis for a three-valued logic on distributions. A word w is hashed to a location ($\chi_S(w) = 2$), a word w is not hashed to a location ($\chi_S(w) = 0$), or ...

Generalization to $\mathbf{n}$ gives an n-valued logic basis with respect to distributions in $(\mathbf{p}^W \times \mathbf{n}^W, +)$ and if we are prepared to use the continuous real line interval, $I = [0, 1]$, as the basis then we are set up for fuzzy set theory and fuzzy logic in $(\mathbf{p}^W \times I^W, +)$.

For the final isomorphic representation of

$$\mathbf{Z}_p \times \mathcal{P}W \equiv \mathbf{p}^W \times \mathbf{2}^W$$

we may employ the law of exponents to write

$$\mathbf{Z}_p \times \mathcal{P}W \equiv \mathbf{p}^W \times \mathbf{2}^W \equiv (\mathbf{p} \times \mathbf{2})^W$$

This new total function space from words $w \in W$ to pairs $(h, b) \in (\mathbf{p} \times \mathbf{2})^W$ can also be expressed in the form of a space of total functions of the **direct sum** of the two abelian groups $\mathbf{Z}_p$ and $\mathbf{Z}_2$

$$\delta \in (\mathbf{Z}_p \oplus \mathbf{Z}_2)^W$$

where we have chosen δ to emphasize that here we are dealing with a space of distributions. That one should cast the model in terms of direct sums lies directly at the heart of the search for a geometry of formal methods. If we use our simple running example, then we can represent δ as

$$\delta = \begin{bmatrix} \vdots & \mapsto & (\_, 0) \\ u & \mapsto & (i, 1) \\ v & \mapsto & (j, 1) \\ w & \mapsto & (k, 1) \\ \vdots & \mapsto & (\_, 0) \end{bmatrix}$$

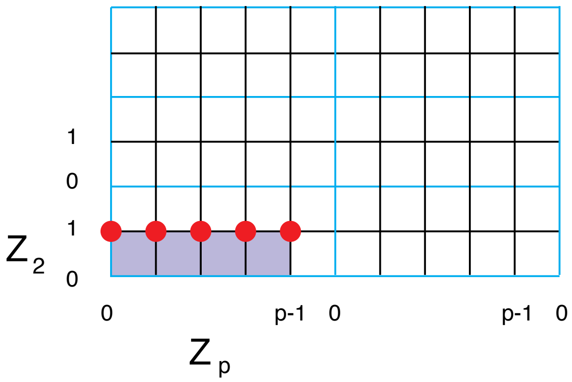

Figure 2: Direct sum $\mathbf{Z}_p \oplus \mathbf{Z}_2$

We may think of $(\mathbf{Z}_p \oplus \mathbf{Z}_2)$ as representing a 2-dimensional discrete space modulo p in the 'X-direction' and modulo 2 in the 'Y-direction'. For convenience, we may exhibit this space as an embedding within the standard euclidean plane and tiling the 2-dimensional plane with a *period parallelogram* as shown (Fig. 2). On the other hand, it is quite natural to wrap this space onto a torus (Fig. 3). Note that only the points marked are part of the discrete space. Such a discrete torus consists of the product of p vertical rings times 2 horizontal rings. Points of the form $(k, 0)$ are shown on the inner horizontal ring. In either case, it is obvious that the natural 'orthonormal' basis is given by $(0, 1)$ and $(1, 0)$. Words which are distributed under δ occupy those positions or sites indicated by the discs on the *outer* horizontal ring.

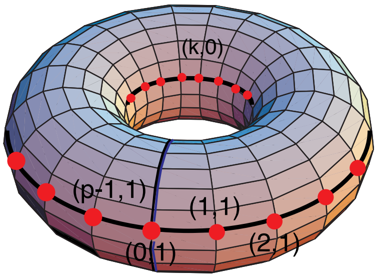

Figure 3: Direct sum $\mathbf{Z}_p \oplus \mathbf{Z}_2$ on the torus

4 Geometry

From a geometrical perspective a hash table may be considered to be a cross-section (abbreviated *section*) η of a fibre bundle, a result which was briefly mentioned in an earlier paper [18, 11–2]. Our intuition at that time was based on Fig. 4 and it seemed abundantly clear that the "fibre over some $j \in \mathbf{Z}_p$, denoted $h^{-1}(j)$, is nothing other than the set of words which hash to j". But when it came to checking details and in particular when care was taken to establish an exact correspondence with the usual definition of fibre bundle, we realized that it would be more appropriate to *invert* the relation and adopt what *appeared* to be a completely counter-intuitive view (at least from the usual perspective of the computer scientist/software engineer)! Lest the informed reader have any doubts, let us briefly remark that the hash function has already been identified as one of the numerous $\Sigma^\star$-morphisms of the free monoid [15]. Consequently, should the *base* space of words prove to be the 'right view', then **all** other $\Sigma^\star$-morphisms immediately become candidates as

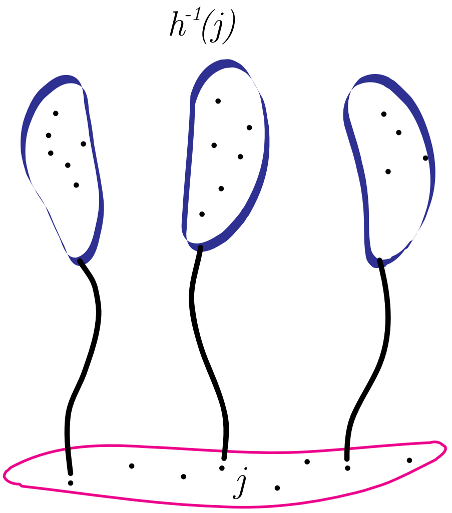

Figure 4: Fibre bundles as asparagus

sections of suitably abstracted/generalized fibre bundles or possibly as bundle morphisms. This would be a major result in the geometry of formal methods!

First let us give a *working definition* of fibre bundle (and here we are cautious to point out that we *do* deliberately omit all considerations of a topological nature at this point but return to the matter at the end of the paper), adopted from Burke [4, 85–7] who, from the perspective of (theoretical) physics, considers (i) the idea of a field to be a simple extension of a function whereby at each point in a given space some geometric object is specified and (ii) to generalize so as to arrange for the comparison of field values at different points. For the very special case with which we are concerned the geometric object is the additive abelian group $\mathbf{Z}_p$ which curiously/fortuitously is also known as a *finite field* (see [14]). This ambiguity in the meaning of *field* will allow us to use the word in the working definition:

Definition 4.1 *A fibre bundle is a pair of spaces, say, E and M, and a projection map from E to M, $\pi\colon E \to M$. The space M is the base space whereon the field is defined. The space of interest, E, is that larger space where we* graph *the field, and the projection π, is a map that assigns to each point on the graph the point in M where that field variable is defined.*

Such a fibre bundle may be denoted simply as

$$\begin{tikzcd} E \arrow[r, "\pi"] & M \end{tikzcd}$$

Our original intuition seems to be entirely in agreement with this definition, in the following sense. The base space M is, of course, the (discrete) space of points $\{0, 1, \ldots, j, \ldots p-1\}$ taken from $\mathbf{Z}_p$ and the space of interest E is the product of the base $\mathbf{Z}_p$ and the space of set of words $\mathcal{P}W$. The projection is naturally the hash function h. We sum up this interpretation with the diagram

$$\begin{tikzcd} \mathbf{Z}_p \times \mathcal{P}W \arrow[r, "h"] & \mathbf{Z}_p \end{tikzcd}$$

Thus, associated with a point j (the hash value) we have a particular set of words $h^{-1}(j)$ (the words which hash to j). The fibre over j, E_j, may be written as $\{j\} \times \mathcal{P}W$. [**Aside:** Technically speaking, all of the fibre bundles that we have unearthed in the specifications of formal methods so far have been *trivial* bundles globally and hence are simply products. We are actively searching for (instances of) bundles which are not globally products. Such a discovery would suggest new directions for computing.]

For a given fibre bundle, the concept of a (cross-)*section* is important. Essentially, from each of the 'heads' of the fibre we select a particular element. Formally, again from Burke [4]

> **Definition 4.2** *A particular field is called a section and is given by a map from the base space to the larger space of interest:*
>
> $$\Gamma : M \to E$$
>
> *such that the projection back down to the base brings us to the point from which we started:*
>
> $$\pi \circ \Gamma(u) = u$$

Applying this notion to our intuitive understanding of the hash table puts us in a quandry. Specifically, from a practical point of view, which is extremely important if not *all important* in applied mathematics, we did not find a reasonable or sensible interpretation originally! Why should a particular collection of p words or fewer have a particular meaning in the context of hashing? Then, one interpretation emerged. Such a small collection of words is directly related to the search for the *perfect hash table* as used for the keywords of a programming language. If we have p distinct keywords then they form the section for the perfect hashing function. If there are just m words, m not a prime, and p is the next prime number larger than m then we can fill the section with null words. Such a fibre bundle may be written as

$$\begin{tikzcd} W \arrow[r] & \mathbf{Z}_p \times W \arrow[r, "h"] & \mathbf{Z}_p \end{tikzcd}$$

where W is the standard fibre. A typical fibre over a page index j has the form (j, u) for some word u which hashes to j under h. [**Aside:** Strictly one would have expected to see $\mathcal{P}W \to \mathbf{Z}_p \times W \xrightarrow{h} \mathbf{Z}_p$. But since elements of the fibres are of the form $(j, \{u\})$ then we may identify singleton words with words and use elements of the form (j, u).] Referring back to Fig. 4 it is clear that our original intuitive understanding of hash table as fibre bundle is really very special.

We will now develop an alternative view of hash table in the next sub-section, one which initially seemed so counter-intuitive, but which proved much more fruitful and exciting.

4.1 Hash table as section of a bundle

Originally, we had chosen $\mathbf{Z}_p$ as the base for the fibre bundle associated with the hash table. It was in the course of trying to determine the meaning of rehashing in such a general context that it occurred to us to take the opposite view and to choose W as the base, where for completeness we consider the words to be elements of a (non-abelian) free group. In this case we have

$$\begin{tikzcd} F \arrow[r] & E \arrow[r, "\pi"] & M \end{tikzcd}$$

where the standard fibre $F = \mathbf{Z}_p$, the space of interest is $E = W \times \mathbf{Z}_p$, and the base is $M = W = FG(\Sigma)$.

For each word u in W the fibre over u, denoted E_u, is nothing other than (u, j) where j is the page number to which u would be hashed under the hashing function.

A *section* η: $M \to E$ associates with each word u in the base space, the pair (u, j), where j is a suitably chosen element of the additive abelian group $\mathbf{Z}_p$. Thus, in this model, a section is the hash function. Consequently, we may restrict this map to a set of words S to give us the (finite) hash table

$$(\eta, S) \in (E^M, \mathcal{P}M) = ((W \times \mathbf{Z}_p)^W, \mathcal{P}W)$$

Comparison with the original algebraic form $(\eta, S) \in (\mathbf{Z}_p^W, \mathcal{P}W)$ given earlier would appear to indicate that there is very close agreement between the two results.

We have proposed a fibre bundle geometry for a hash table (and by extension for distributions). What use can we make of it? Does it assist us better to understand the specification of distributions? In what way does it complement the algebra? We shall explore the answers to these questions now.

Consider first the constant **zero section** $\eta \colon u \mapsto (u, 0)$. It turns out to be of particular practical interest! Such a constant hash function which maps all words to the same location has been considered convenient "when debugging a program" [13, 513]. It also has theoretical significance (see below).

Now let us turn our attention to the composition of sections $\eta * \kappa$ where the binary operator $*$ is to be determined. We expect this to be given in *pointwise* fashion over the space of words

$$\begin{aligned}(\eta * \kappa)(u) &= \eta(u) * \kappa(u)\\ &= (u, j) * (u, k)\\ &= (u, f(j, k))\end{aligned}$$

where we have chosen $f(j, k)$ to denote the most general combination of pages j and k. In the earlier part of the paper on the algebraic aspects of distributions we were guided by the fact that $\mathbf{Z}_p$ was an abelian group and were somehow 'forced' to define

$$f(j, k) = j + k \pmod p$$

From the geometrical perspective it is clear that there are many different possibilities. The important aspect is that whatever the decision of the choice of $f(i, j)$ we must get a fibre over the word u. In other words the combination of different sections can only refer to combinations in the fields and not to the base! This clears up one of the misunderstandings that arose in the algebra. Note also that the operator $*$ is determined by the function f on field elements j and k.

With respect to the application of hashing one reasonable choice for $f(j, k)$ is to choose the override operator. Therefore, $\eta \dagger \kappa$ forces the choice of

$$f(i, j) = j$$

Now if we start with the zero section, denoted η_0, then we can gradually build up a desired hash function and hash table by repeated overrides. There may be other interesting possibilities for distributions other than hashing, and even for hashing itself! It is worth noting that Knuth [13, 521] refers to the use of a pair of hash functions, say (η, κ), in order to implement *hashing using open addressing*. This is used to avoid the overflow chaining mechanism which we have assumed. This practical hashing algorithm may be set in the following geometric context of the fibre bundle. Given sections η and κ. What is the relationship between $\eta(u)$ and $\kappa(v)$ for all words u and v? How might such pairs (η, κ) be chosen?

Finally let us consider expressions of the form $\eta(u) * \eta(v)$ where we are interested in the same section η. What might be the nature of the operation? Expanding out we obtain

$$\begin{aligned}\eta(u) * \eta(v) &= (u, j) * (v, k)\\ &= (u * v, j * k)\end{aligned}$$

Could the resulting page index $j * k$ be related to the structure of the word space and in particular to the composition $u * v$? Suppose that $u * v$ denoted the concatenation of u and v. What would $j * k$ correspond to?

These questions prompt us to explore the possibility of determining trajectories which might correspond to incremental changes in words (Fig. 5). Suppose that we started with a word $u =$ '*art*' and by concatenation passed through the points $u_1 =$ '*artist*', $u_2 =$ '*artistic*', ending up at $v =$ '*artistically*'. It is intuitively clear that we traverse a curve or trajectory in the base space of words. Might there not be a comparable trajectory in the fibre space? Does not such a trajectory constitute a section? Might one not therefore construct a hash function that takes into account such smoothness?

The questions are by their very nature extremely interesting from both a geometrical perspective—the geometry of computing—and from an algorithmic perspective—algorithms/programs are paths in an

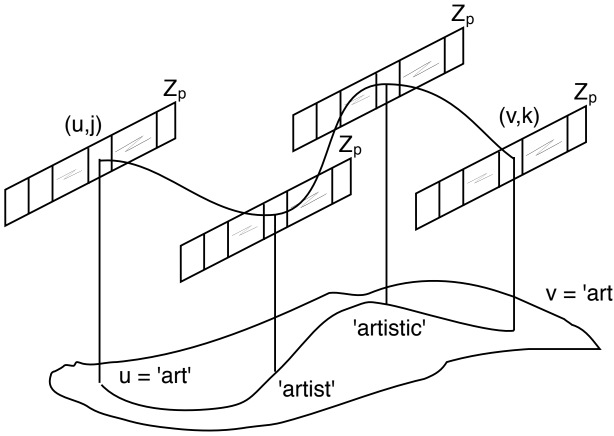

Figure 5: Trajectories

appropriate geometric computation space. With respect to the specific problem domain of hashing it is noteworthy that Knuth has already referred to the importance of the *multiplicative hash method* in taking advantage of nonrandomness of words [13, 510]. Admittedly, he cites clusters such as {TYPEA, TYPEB, TYPEC} which are of a different nature to those on the trajectory. But the principle is there and is worth exploring. The other important aspect is clearly the notion of smooth curve or trajectory, both in the word space and in the bundle space. Smoothness calls to mind continuity (of varying orders or degrees). Later in the paper we will show that there is indeed first order continuity of curves in the word space.

4.2 Rehashing as bundle morphism

To test further the appropriateness of this geometrical approach let us revisit the practical phenomenon of *rehashing* introduced earlier. We will show that rehashing may be considered to be a fibre bundle morphism [5, 122] (Fig. 6)

$$\Phi(u, j) = (\varphi(u), g_u j)$$

where g_u is a morphism from the additive abelian group $\mathbf{Z}_p$ to $\mathbf{Z}_q$. [**Aside:** In computing, g_u might be written as $g[\![u]\!]$ to emphasize that it is a curried function or operator.] Note that the diagram (Fig. 6) has

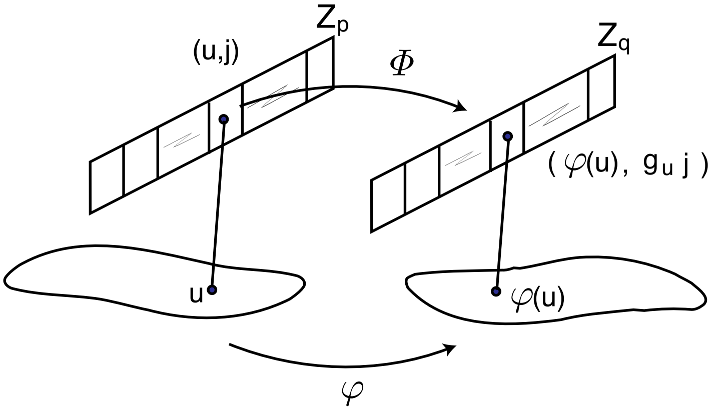

Figure 6: Bundle morphism

the form of a commuting diagram that would be familiar to anyone who conducts reification proofs in formal methods. This may be made more explicit by introducing sections Γ and Γ'. Then the bundle morphism may be expressed in terms of

$$(\Phi \circ \Gamma)(u) = (\Gamma' \circ \phi)(u')$$

Since we require that the same words be rehashed then we will naturally choose φ to be the identity function on words. In this case the bundle morphism takes the form

$$\Phi(u, j) = (u, g_u j)$$

Consequently it is easy to see that everything depends on the group morphism $g_u \colon \mathbf{Z}_p \to \mathbf{Z}_q$. For rehashing we require that q be somewhat bigger than p. If we adhere to this relationship then the group morphism g can not possibily cover all elements of $\mathbf{Z}_q$. It must be an inclusion or a permutation of an inclusion. If we want to locate rehashing within this structural scheme then we interpret the group morphism in the *opposite sense*, a morphism from the already *rehashed* structure $\mathbf{Z}_q$ to the original $\mathbf{Z}_p$, in which case the morphism arrows of Fig 6 will be reversed.

For g_u to be a *group* morphism, we must have the relations that (1) the image of the identity in $\mathbf{Z}_p$ is the identity in $\mathbf{Z}_q$ and (2) if k is the image of an element j then their inverses must also match up:

$$g_u 0 = 0, \qquad g_u j = k, \qquad g_u(p - j) = q - k$$

Rehashing is clearly a morphism of sections, i.e., of hash tables. This is the result we seek. It remains for us to investigate if there are practical consequences of this geometrical view.

4.3 Parikh mapping

Let us consider another example of a fibre bundle constructed on the base of words W. In this case we will use the Parikh mapping ψ which returns a vector of the counts of the letters in a given word. For definiteness we will consider the alphabet $\Sigma = \{a, b, c\}$ and extend the words to elements of the corresponding free group. Consider, for example, the application of mapping ψ to a few words for illustration:

$$\psi(abc) = (1, 1, 1), \qquad \psi(abac) = (2, 1, 1), \qquad \psi(ab\bar{a}c) = (0, 1, 1)$$

For the fibre we will choose the direct sum $F = \mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z}$ and the basis elements are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ which correspond to the letters a, b, and c, respectively.

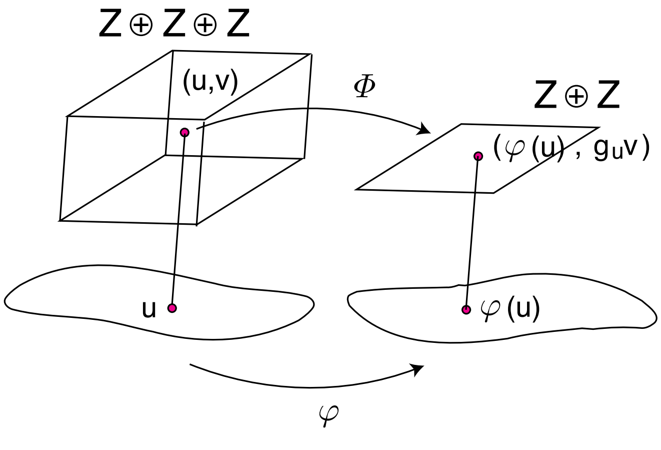

Figure 7: Parikh mapping

$$\begin{tikzcd}(W \to \mathcal{P}D) \arrow[r, shift left, "\mathcal{D}_{\mathcal{S}}"] & (L \to (W \to \mathcal{P}D)) \arrow[l, shift left, "\mathcal{R}"]\end{tikzcd}$$

Figure 8: Distribution and retrieval

Let us now introduce a bundle morphism by specifying $\varphi = ⩤_a$. Since we are working in the free group, we intend $⩤_a$ to have the meaning of the removal of *both a and* its formal inverse $\bar{a}$. In these circumstances, the morphism is

$$\Phi(u, \mathbf{v}) = (\underset{a}{⩤}\, u, g_u\mathbf{v})$$

The group element $g_u \colon \mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z} \to \mathbf{Z} \oplus \mathbf{Z}$ forgets the first coordinate $g_u(i, j, k) = (j, k)$ and consequently is a projection mapping.

This fibre bundle turns out to be a paradigmatically useful geometrical model for a variety of specifications including the distribution of the dictionary introduced in the prologue. It is paradigmatic in the sense that we might consider a location, node, etc., as being represented by an element in a direct sum space of dimension n, ${}^{\oplus}/(\mathbf{Z})_i$, $1 \le i \le n$. For example, if we wish to model a *crossword* then a fibre bundle is a natural choice. The base is W, the space of words. For the standard fibre we choose $\mathbf{Z} \oplus \mathbf{Z} \oplus \mathbf{Z}$. The section $\Gamma \colon u \mapsto (u, (i, j, k))$ gives the crossword where the word u is to be filled in at coordinates (i, j) and has length k.

Now let us determine the fibre bundle for the spelling-checker dictionary introduced in the Prologue. In the light of the previous example, it is clear that we ought to consider the dictionary to be given by a section of the bundle

$$\begin{tikzcd}L \arrow[r] & W \times L \arrow[r, "\pi"] & W\end{tikzcd}$$

where we define the section $\Gamma \colon W \to W \times L$ by $\Gamma \colon u \mapsto (u, l)$. Clearly this corresponds exactly to $A_{l,u} \colon \delta \mapsto \delta \textcircled{$\cup$} [u \mapsto \{l\}]$. In addition, the latter expression is the algebraic equivalent of the intersection of the section and the 'level curve' for location l.

To complete this introduction to the application of fibre bundles to specifications let us consider a slightly more complex version of the distributed dictionary introduced in the Prologue. Suppose that we start with a simple dictionary modelled as the space of maps from words to their definitions

$$\delta \in \mathit{DICT} = W \to \mathcal{P}D$$

Now we distribute the dictionary over several locations such that there is no overlap or sharing of words and their definitions. In short, the resulting distributed dictionary is to be a partition of the original. We may model this distributed dictionary as the space of maps from locations to dictionaries

$$\delta \in \mathit{DICT} = L \to (W \to \mathcal{P}D)$$

subject to an appropriate invariant and retrieval function. We propose to **define this distribution completely** by a fibre bundle and hence we refuse to pre-judge the form of the retrieval function $\mathcal{R}$ as shown in Fig. 8. For the standard fibre $F = L$ it is obvious that we ought to choose a direct sum space $F = {}^{\oplus}/(\mathbf{Z})_i$ to stand for any reasonable space which will model locations (and this would include both hardware addresses as well as WWW URLs). For the base we will choose the space of dictionaries. Hence with each dictionary δ_j we will associate a location j. From the geometry, it is clear that we ought to define the distributed dictionary as a section $\Gamma \colon M \to M \times E$ such that

$$\Gamma(\delta_j) = (\delta_j, j)$$

where j ranges over locations. Based on our analysis of the hash table earlier we look at (i) the composition of distinct sections $\Gamma_1 * \Gamma_2$ which are defined pointwise over the base space

$$(\Gamma_1 * \Gamma_2)(\delta) = \Gamma_1(\delta) * \Gamma_2(\delta)$$

and (ii) compositions of the form $\Gamma(\delta) * \Gamma(\delta')$, where the operator $*$ is yet to be determined. Expansion of the pointwise composition gives

$$\begin{aligned}(\Gamma_1 * \Gamma_2)(\delta) &= \Gamma_1(\delta) * \Gamma_2(\delta)\\ &= (\delta, l_1) * (\delta, l_2)\\ &= (\delta, f(l_1, l_2))\end{aligned}$$

where we use $f(l_1, l_2)$ again to denote a general composition on locations. As for the case of the hash table the only practical meaningful choice would appear to be the override operator. That does not mean that there are no others . . . We just have not been able to find them.

The second composition $\Gamma(\delta) * \Gamma(\delta')$ leads to

$$\begin{aligned}\Gamma(\delta) * \Gamma(\delta') &= (\delta, l_1) * (\delta', l_2)\\ &= (\delta * \delta', l_1 * l_2)\end{aligned}$$

By the assumption that the distribution is a strict partition, then we have $\mathsf{dom}\,\delta \cap \mathsf{dom}\,\delta' = \emptyset$ and the composition may be written

$$\Gamma(\delta) * \Gamma(\delta') = (\delta \sqcup \delta', f(l_1, l_2))$$

But there are other ways in which to combine dictionaries. In the beginning of the paper we saw that a 'natural' algebraic structure was obtained if the space of dictionaries was an indexed monoid. Thus for two dictionaries δ and δ', we may combine them to obtain $\delta \mathbin{\bigcirc\kern-0.75em\cup} \delta'$. In the geometry the combination is to be on the sections $\Gamma(\delta)$ and $\Gamma(\delta')$. At present we can only speculate that perhaps if we give up the idea of strict partitioning then we can find a bundle endomorphism such that the composition $\Gamma(\delta) * \Gamma(\delta')$ is directly related to the composition of dictionaries $\delta \mathbin{\bigcirc\kern-0.75em\cup} \delta'$. Finally, we would expect that in analogy to the spelling checker dictionary the intersection of the section and the level curve for a location l would lead to an algebraic expression of the form $\kappa \mathbin{\bigcirc\kern-0.75em\cup}^2 [l \mapsto \delta]$, where

$$\kappa \mathbin{\bigcirc\kern-0.75em\cup}^2 [l \mapsto \delta] = \begin{cases} \kappa \sqcup [l \mapsto \delta], & \text{if } \neg\chi_\kappa(l)\\ \kappa \dagger [l \mapsto \kappa(l) \mathbin{\bigcirc\kern-0.75em\cup} \delta], & \text{otherwise}\end{cases}$$

4.4 Topological considerations

There is more to fibre bundles [19]. All of the bundles we have seen so far are *trivial* bundles. They are nothing other than the cartesian products of a pair of spaces, the base space and the fibre. To illustrate that more is involved in the concept of a fibre bundle we will pick a classical illustration from geometry (see Fig. 9) Both the cylinder and the Möbius band are constructed as ruled surfaces, i.e., they are nothing more than collections of straight lines (i.e., bundles of fibres). The cylinder is generated by the quaternion equation

$$S(u, q) = qL_u(a, b)q^{-1}$$

where $q = exp(n\theta/2)$ and the generator line $L_u(a, b) = (1 - u)a + ub$, $u \in [0, 1]$, is rotated about an axis with direction vector $n = (b - a)/|b - a|$. The point a moves on the circle. As a fibre bundle it is completely given by

$$\begin{tikzcd}\mathbf{R} \arrow[r] & S^1 \times \mathbf{R} \arrow[r, "\pi"] & S^1\end{tikzcd}$$

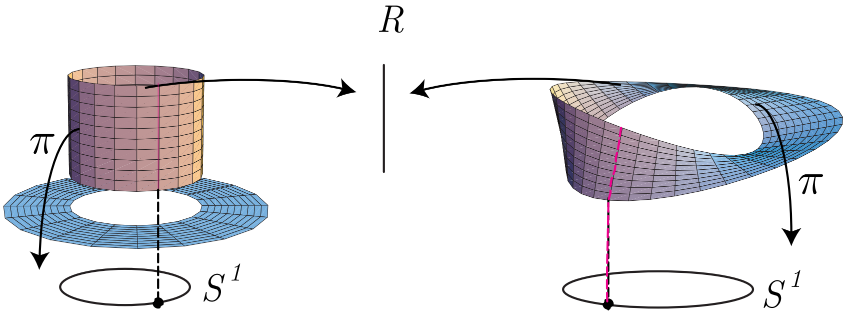

Figure 9: Cylinder and Möbius band

The fibre bundle may clearly be regarded as a *decomposition* of the cylinder. In the case of the Möbius band the quaternion equation of which has the form

$$S(u,q) = q\big(e^{\varphi t}(L_u(a,b) - m)\big)q^{-1} + qmq^{-1}$$

it is the midpoint m of the generator line $L_u(a,b)$ which moves on the circle. The generator line rotates about a central axis n **and** about a tangent vector t to the circle at the same time. The fibre bundle is different from that of the cylinder:

$$\begin{tikzcd} \mathbf{R} \arrow[r] & E \arrow[r, "\pi"] & S^1 \end{tikzcd}$$

where we have used E to denote the bundle space in question. The difference between the two structures is explained by the *twist* in the Möbius band. Topologically, the difference may be described by *transition functions*. Hitherto all of our bundles are trivial and hence the group of transition functions is the trivial group of the identity. It is an open question whether there are non-trivial bundles of interest to specifications in formal methods. There is one other point to be resolved.

In order that we might use the term *fibre bundle* validly to describe this geometry we need to be able to demonstrate that our spaces are manifolds in some real sense. We need to produce open sets and hence demonstrate continuity of functions. Relevant technical details are relegated to the appendix.

Let the base $W = (\Sigma \cup \bar{\Sigma})^{\mathbf{N}}$ be equipped with the *finite topology* of function spaces. The open sets $O_{u,\mathbf{k}}$ are indexed with words u in $(\Sigma \cup \bar{\Sigma})^{\mathbf{N}}$ and consists of all those words v which share the same initial segment (i.e., prefix) of length k, denoted in Eilenberg's notation $u^{[k]} = v^{[k]}$ [6, 358]. Next take the standard fibre F. Equip it at least with the discrete topology. The product manifold $E = W \times F$ may now be given the usual product set topology. The fibres over u, denoted E_u, are contained in open sets $\pi^{-1}(O_{u,\mathbf{k}})$ and the sections $\Gamma\colon W \to E$ are continuous. This sets the scene for a fibre bundle geometry over words, lists, sequences.

But more is possible. The finite topology of function spaces Y^X is applicable also to subsets of the total function space (via the induced topology) and hence to the usual domains of maps (i.e., partial functions) in VDM and Z. Therefore, the base space of the distributed dictionary, $M = W \;\to\; \mathcal{P}D$, may also be given the finite set topology and hence the corresponding sections, i.e., distributions, are continuous. We believe that with this paper we have opened the door on an exciting new world of geometric algebra, the geometry of formal specifications.

5 Epilogue

> "*Both [Abel and Jacobi] had arrived at the key idea of working with inverse functions of the elliptic integrals [and thus discovered the beautiful domain of elliptic functions]*" [12, 646]

There is a geometry of specifications. A first account of initial results was reported at the *Northern Formal Methods Workshop* in Ilkley, UK, 1996 [18]. Then we demonstrated that tail-recursive **functions/algorithms** correspond to affine transformations. Here we present for the first time the setting of **data structures** and by extension the spaces of distribution models in a fibre bundle context.

The *discovery of the geometry of fibre bundles in specifications* is similar to that of the discovery of elliptic functions *in the sense* that it was only possible by using the **principle of inversion** [17]. By this we mean that it was necessary to invert the way in which we thought about hash functions and hash tables in order to obtain our results. Now it all seems so obvious. In 1996 it was not so.

There are many directions now opened up for future development. Outstanding is the search for non-trivial bundles of data structures. Such bundles can probably be *constructed*. Whether they prove to be of use in practical implementations is not yet known.

The section in the paper on isomorphic representations which terminated with the use of the torus for the embedding of the hash function clearly indicates that there are probably many points of contact between the geometric algebra being developed and 'geometry in the classical sense'.

Finally, it behoves us make a remark or two upon our choice of fibre bundles over sheaves. In a more general setting, it is clear that sheaves (and Sheaf Theory) offer much more potential for the kind of geometry that we seek. The work by Goguen, for example, on the semantics of concurrent (object-oriented) systems [7] shows what is possible. However, we have always felt that we understand exactly how engineers really think and we are convinced that in the *first instance* any introduction of fibre bundle will be much more acceptable than sheaf. On the other hand, we recognize that we were very fortunate to choose an example—the hash table—that permitted us to introduce some of the key aspects of fibre bundles as distinct from sheaves: the group actions and (finite) fields. In order that sheaf theory be seen to be of *practical benefit* then practical examples such as that given on hashing and distributions must be exhibited/constructed. This is clearly work for another day.

5.1 Acknowledgements

Thanks goes to all the members of the *Formal Methods Working Group* on the one hand and the students on the other, at the University of Dublin, Trinity College, who have been patient and tolerant enough to have endured much of the development described in the paper.

In an earlier version of this paper we had thought to confine the geometry to fibre bundles. However, the thoughtful comments of one of the anonomous referees and a subsequent conversation with Dr. Tony Seda (Department of Mathematics, University College Cork) convinced us that we ought to include also some salient remarks on the relevance of sheaf theory to our work. Indeed, an account of distributions in terms of sheaves is currently under way.

As is customary, the principal author acknowledges that any errors of whatever form that remain in the paper are his sole responsibility.

The commuting diagrams in this paper were produced by the TEX package diagram.tex[1] of Paul Taylor, Department of Computer Science, Queen Mary and Westfield College, London E1 4NS, UK (e-mail: pt@dcs.qmw.ac.uk).

The bibliography is produced with the aid of MacBibTEX[2] of Vince Darley, Division of Applied Sciences, Harvard University, Oxford Street, Cambridge MA 02138, USA (e-mail: vince@das.harvard.edu).

The torus, cylinder, annulus, and Möbius band were all generated by *Mathematica 3.0* [20] directly from their quaternion equations.

[1] ftp://ftp.dcs.qmw.ac.uk/pub/tex/contrib/pt/diagrams/

[2] http://www.fas.harvard.edu/˜darley/

6 Appendix

6.1 Hash table specification

For convenience we present an outline of an abstract model of a hash table as a sequence of p pages, p a prime, each page containing a sorted sequence of words:

$$\begin{aligned} \delta \in \text{HASH\_TABLE} &= \text{PAGE}^p \\ \text{PAGE} &= W^\star \end{aligned}$$

subject to an appropriate invariant:

$$\text{inv-PAGE}(\delta_{\eta(w)}) \stackrel{\triangle}{=} (\sigma(\delta_{\eta(w)}) = \delta_{\eta(w)}) \wedge (|\,\mathsf{elems}\,\delta_{\eta(w)}| = |\,\mathsf{len}\,\delta_{\eta(w)}|)$$

where σ denotes the sort morphism. That a given word can only appear on one page is deemed to be a property of the hash function. A hash function is an element of

$$\eta \in \text{HASH\_FUNC} = W \to \mathbf{Z}_p$$

The addition of a *new* word w is given by

$$A\colon \text{HASH\_FUNC} \times W \longrightarrow \text{HASH\_TABLE} \longrightarrow \text{HASH\_TABLE}$$

$$A(\eta, w)(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r) \stackrel{\triangle}{=} (\delta_l \cdot \langle\, \sigma(\langle\, w \,\rangle \cdot \delta_{\eta(w)}) \,\rangle \cdot \delta_r)$$

where $\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r$ denotes the partitioning of a hash table δ uniquely into left δ_l, middle $\langle\, \delta_{\eta(w)} \,\rangle$ and right δ_r parts. The concept of *newness* is guaranteed by the pre-condition

$$\text{pre-}A(\eta, w)(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r) \stackrel{\triangle}{=} \neg\chi_{\delta(\eta(w))}(w)$$

Removal of an existing word w is specified by

$$R\colon \text{HASH\_FUNC} \times W \longrightarrow \text{HASH\_TABLE} \longrightarrow \text{HASH\_TABLE}$$

$$R(\eta, w)(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r) \stackrel{\triangle}{=} (\delta_l \cdot \langle\, \lhd\!\!\!-_{\,w}\, \delta_{\eta(w)} \,\rangle \cdot \delta_r)$$

That the word w already *exists* in the hash table requires a pre-condition.
To look up an existing word w is specified by

$$L\colon \text{HASH\_FUNC} \times W \longrightarrow \text{HASH\_TABLE} \longrightarrow \mathbf{B}$$

$$L(\eta, w)(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r) \stackrel{\triangle}{=} \chi_{\delta(\eta(w))}(w)$$

[**Aside:** Note that we have used $A(\eta, w)(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r)$ instead of $A_{\eta,w}(\delta_l \cdot \langle\, \delta_{\eta(w)} \,\rangle \cdot \delta_r)$, for clarity.]

6.2 Topology

In order to be able to claim that we do have a geometry of formal methods we must at least give some indication that there is a sensible topology associated with our spaces.

Discrete topology Let us begin with a simple definition [1, 28].

> **Definition 6.1** *A topology on a set X is a nonempty collection of subsets of X, called* open sets, *such that any union of open sets is open, any* finite *intersection of open sets is open, and both X and the empty set are open. A set together with a topology on it is called a topological space.*

For each set X we may use $\mathcal{P}X$ to form the *discrete topology* on X. Specifically, every element in $\mathcal{P}X$ is an open set. This is the largest possible topology on X. With this topology any function with domain W is continuous [1, 14].

Consider the set W of words. Let W be equipped with the *discrete topology.*

Subspace topology Given a topological space X and a subset Y of X. Form the intersection of all the open subsets of X with Y to obtain the *subspace* or induced topology on Y [1, 28].

Let S be a subset of W, the space of words furnished with the discrete topology. Then the subspace topology on S is the induced discrete topology. The open sets are elements of $\mathcal{P}S$.

Basic open sets Given a topology on a set X, and a collection B of open sets such that every open set is a union of members of B. Then S is called a *base* for the topology and elements of B are called *basic open sets* [1, 30].

Consider the discrete topological space W. Define B to be the set which consists of the empty set and all the singleton sets of $\mathcal{P}W$. Then B is a base for the discrete topology. Since the intersection of any pair of elements in B is the empty set then we say that B forms a discrete base.

Continuous functions Let X and Y be topological spaces. A function from X to Y is continuous if and only if the inverse image of each open set of Y is open in X [1, 32]. One way in which to determine if a function is continuous is to investigate the inverse images of the elements of the base B for Y. Let f be a function from X to Y. If B is a base for Y and the inverse image of each element of the base is open in X then f is continuous.

Product topology Let X and Y be topological spaces and let B denote the family of all subsets of $X \times Y$ of the form $U \times V$, where U is open in X and V is open in Y. Then $^{\cup}/B = X \times Y$ and the intersection of any two members of B lies in B. Therefore B is a base for a topology on $X \times Y$. This topology is called the *product topology* [1, 52].

The functions $p_1 \colon X \times Y \to X$ and $p_2 \colon X \times Y \to Y$ defined by $p_1(x, y) = x$ and $p_2(x, y) = y$ are called projections. The projection functions are continuous.

Finite topology for function spaces Let X and Y be arbitrary sets for which Y has the discrete topology. Then the set of maps Y^X can be given the product topology formed out of $|X|$ copies of Y [9, 469]. A base for the open sets of this topology, called the *finite topology* for Y^X, consists of the sets $O_{f,\{x_i\}}$ where f is a map $X \to Y$, $\{x_i \mid 1 \le i \le n\}$ a finite subset of X, and

$$O_{f,\{x_i\}} = \{g \in Y^X \mid g(x_i) = f(x_i), 1 \le i \le n\}$$

Now let us choose $X = \mathbf{N}$ and $Y = \Sigma$. The corresponding function space $Y^X = \Sigma^{\mathbf{N}}$ consists of (potentially infinite) words over the alphabet Σ. We choose to extend this to include formal letter inverses: $(\Sigma \cup \bar{\Sigma})^{\mathbf{N}}$. Let $u \in (\Sigma \cup \bar{\Sigma})^{\mathbf{N}}$ denote a typical word and $\mathbf{k} = \{1, 2, \dots, k\}$ the set of the first k natural numbers. The open sets, indexed by u, are all the words v which have a common prefix with u,

$$O_{u,\mathbf{k}} = \{v \in (\Sigma \cup \bar{\Sigma})^{\mathbf{N}} \mid v_i = u_i, 1 \le i \le k\}$$

If we denote by W this topological word space, then it may be shown that W is compact [6, 361] with a distance function δ given by

$$\delta(u, v) = 1/n \quad \text{if } u^{[n]} \neq v^{[n]}, \quad u^{[n-1]} = v^{[n-1]}$$

where $u^{[n]}$ denotes the initial segment of length n of u. The distance function turns $(\Sigma \cup \bar{\Sigma})^{\mathbf{N}}$ into a metric space. Moreover, and most importantly for the fibre bundle treatment of the paper, for a *fixed* $u \in \Sigma \cup \bar{\Sigma}^{\star}$, the function $v \to uv$ is continuous (details in [6, 359]).

7 Glossary

$\emptyset$, θ	empty set, empty map
$\{a\}$, $[x \mapsto y]$	singleton set, singleton map
$\mathcal{P}X$	powerset of X
$X \;\rightarrow Y$	the space of all (partial) functions from X to Y
Y^X	th e space of all total functions from X to Y; $Y^X \subset X \;\rightarrow Y$
$U \bigtriangleup V$	symmetric difference of two sets
$f \circ g$	f after g
${}^{\cup}/$, $\bigcup$	reduction with respect to set union
$\delta_j \sqcup \delta_k$	map extension, disjoint union of two maps; defined if $\mathsf{dom}\,\delta_j \cap \mathsf{dom}\,\delta_k = \emptyset$
$\delta_j \circledcup \delta_k$	relational union
$\delta_j \dagger \delta_k$	override
$\chi_\delta(l), \chi[\![l]\!]\delta$	characteristic function or subobject classifier, l in $\mathsf{dom}\,\delta$
$\lhd_U(\eta)$,$\eta\vert_U$	restriction of η to U
$\mathrel{\lhd\!\!\!-}_U(\eta)$, $\eta\backslash U$	removal of η by U
$A_{l,u}, A[\![l, u]\!]$	curried function with args l and u

References

[1] M. A. Armstrong. *Basic Topology.* McGraw-Hill Book Company (UK), London, 1979.

[2] June Barrow-Green. *Poincaré and the Three Body Problem.* Number 11 in History of Mathematics. American Mathematical Society, Providence, Rhode Island, 1997.

[3] J. L. Bell. *Toposes and Local Set Theories, an Introduction.* Number 14 in Oxford Logic Guides. Clarendon Press, Oxford, 1988.

[4] Burke. *Applied Differential Geometry.* Cambridge University Press, Cambridge, 1985.

[5] R. W. R. Darling. *Differential Forms and Connections.* Cambridge University Press, Cambridge, 1994.

[6] Samuel Eilenberg. *Automata, Languages, and Machines, Volume A.* Academic Press, New York, 1974.

[7] Joseph A. Goguen. Sheaf semantics for concurrent interacting objects. *Mathematical Structures in Computer Science*, 2(2):159–191, June 1992.

[8] Robert Goldblatt. *Topoi, The Categorial Analysis of Logic.* North-Holland, Amsterdam, revised edition, [1979] 1984.

[9] Nathan Jacobson. *Basic Algebra II.* W. H. Freeman, New York, second edition, 1989.

[10] P. T. Johnstone. *Topos Theory.* Academic Press, London, 1977.

[11] Victor Kac. *Vertex Algebras for Beginners.* Number 10 in University Lecture Series. American Mathematical Society, Providence, Rhode Island, 1997.

[12] Morris Kline. *Mathematical Thought from Ancient to Modern Times.* Oxford University Press, Oxford, 1972. The 1990 three volume paperback work is cited.

[13] Donald E. Knuth. *The Art of Computer Programming, Vol. 3 Sorting and Searching.* Addison-Wesley, Reading, Massachusetts, 1973.

[14] Rudolf Lidl and Harald Niederreiter. *Introduction to finite fields and their applications.* Cambridge University Press, Cambridge, revised edition, [1986] 1994.

[15] Mícheál Mac an Airchinnigh. *Ph.D. Thesis: Conceptual Models and Computing.* Department of Computer Science, University of Dublin, Trinity College, Dublin, Ireland, 1990.

[16] Mícheál Mac an Airchinnigh. Formal Methods & Testing. In *Tutorials of the Sixth International Software Quality Week*, 625 Third Street, San Francisco, CA 94107–1997, May 1993. Software Research Institute.

[17] Mícheál Mac an Airchinnigh. The Creation of the World in the Third Millenium of our Era. In *Proceedings of Information Technologies and Programming, IT&P'96, Plovdiv*, pages 21–33, Sofia, June 1996. Bulgarian Academy of Sciences.

[18] Mícheál Mac an Airchinnigh. Towards a New Conceptual Framework for the Modelling of Dynamically Distributed Systems. In David Duke and Andy Evans, editors, *Northern Formal Methods Workshop, Ilkley 1996*, Electronic Workshops in Computing. Springer-Verlag, London, 1997. http://www.springer.co.uk/ewic/workshops/.

[19] Walter A. Poor. *Differential Geometric Structures.* McGraw-Hill Book Company, New York, 1981.

[20] Wolfram Research, Inc. *Mathematica.* Wolfram Research, Inc., Champaign, Urbana, 3.0 edition, 1996. The Program.